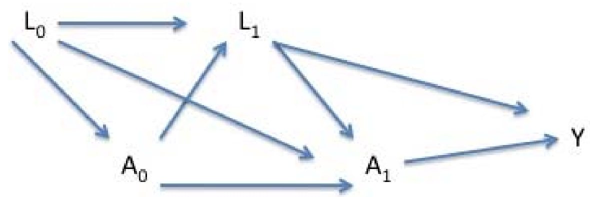
eAppendix1: Illustration of IPTW weight estimation

In the case of non-randomized exposure experienced at a single time point, it is possible to adjust for confounding variables by including them in a regression model. However, when a non-randomized exposure is experienced at multiple time points, covariate adjustment will not work. In the case addressed in this study, imagine the following scenario:



Imagine that L_k denotes a series of confounding variables that exist at time *k* (in our case, this would be, for example, individual income, education and occupation) and A_k denotes the exposure of interest at time *k* (i.e. proportion of neighborhood residents with a family income under the poverty threshold). L_1 is affected by exposure A_0—for example, living in a high-poverty neighborhood may limit the type of income-generating and educational opportunities a person can obtain. At the same time, L_1 confounds the relationship between A_1 and Y—that is, individual income, education and occupation influence the level of exposure to a neighborhood with a certain poverty level, and they are also associated with the alcohol use and misuse. In traditional covariate adjustment, if one adjusts for both A_0 and L_1, one is “overadjusting” for a variable in the causal pathway, thus taking away variability associated with the time-varying treatment. However, if one doesn’t control for L_1, one ignores potential confounding bias.

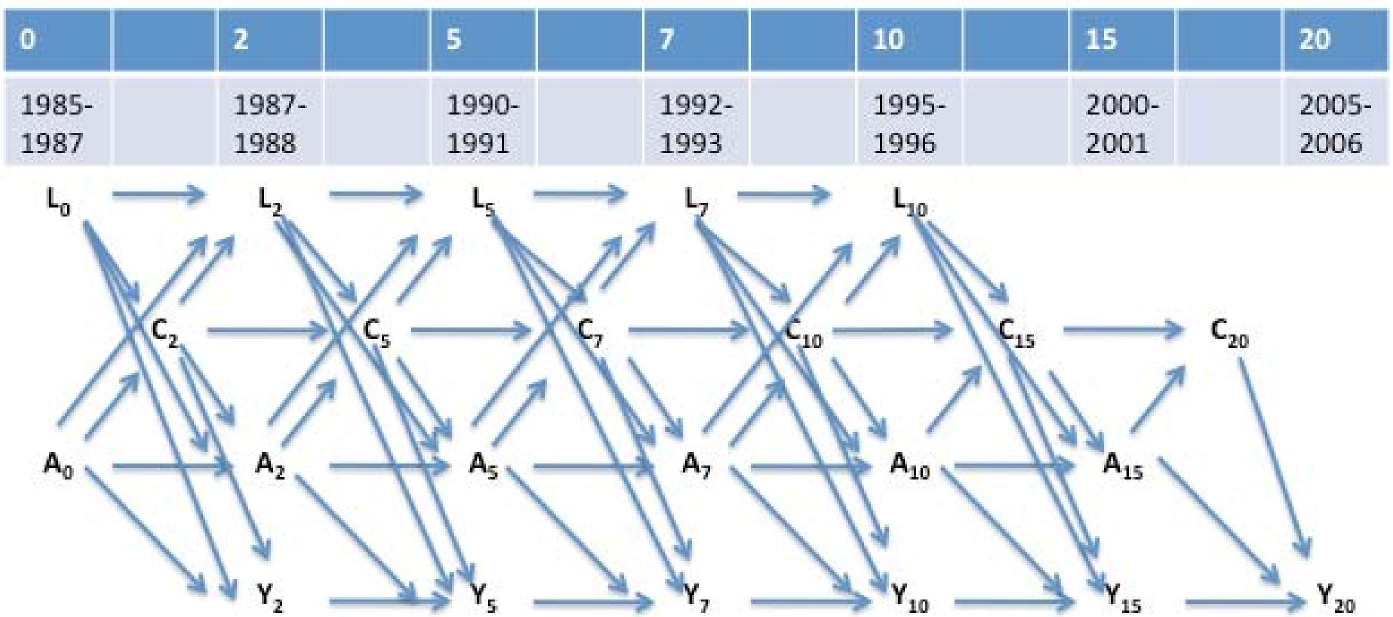
The marginal structural model (MSM) is a tool that can be used in the case of time dependent treatments and time-dependent confounders—i.e. observed covariates that are affected by the treatment and relevant to the outcome of interest. MSMs are estimated using Inverse Probability of Treatment Weighting (IPTW). IPTW calculates the probability of an individual receiving the treatment (exposure in a nonrandomized study) they actually received, conditional on their observed stable and time-varying covariates. Individuals are weighted by the inverse of their probability in order to create a "pseudopopulation" consisting of w_i copies of each subject. People who are most unrepresented in treatment assignment (exposure in a nonrandomized study) are given proportionally higher weights, while individuals who are highly represented in treatment assignment are given proportionately lower weights, so that we can obtain a comparable population in terms of stable and time-varying confounders across levels of the treatment assignment. We can then use the weighted "pseudopopulation" that is balanced in terms of distribution of potential confounders across treatment levels, to estimate the unconfounded relationship between exposure A and outcome Y. By using weighting to address confounding, this approach literally removes time-varying confounders that are in the pathway between the exposure of interest and the outcome, from the dependent side of the equation, and thus avoids the problem of potentially "overcontrolling" for a mediator.

For example, imagine that the distribution of exposure A is imbalanced across the confounder L, so that at L=0, ¾ of the subjects are unexposed to A (A=0) and ¼ are exposed to A (A=1), while at L=1, ¾ of the subjects are exposed to A (A=1) and ¼ are unexposed. If we have 8 subjects (4 at each level of L), and we calculate the probability of A given L, we can conclude that at L=0, 3 of the subjects will have a probability of ¾ of having the exposure A they already have (P (A=0)) and one will have a probability of ¼ of having their own exposure (P (A=1)); in contrast, at L=1, 3 of the subjects will have a probability of ¾ of having A=1, and one will have a probability of ¼ of A=0. Since the IPTW is the inverse of the probability of receiving the treatment they received, given their own covariate history—for those who had L=0 and A=0, the IPTW would be 4/3, while for those who had L=0 and A=1, the IPTW will be 4/1. If, for ease of interpretation, we multiply each of these IPTWs by the relative ratio of the weights (3 to 1), this means that at L=0, the weight for those with A=0 will be 4 and for those with A=1 it will be 12. Using these weights, we will then make 4 copies of each of the three individuals with A=0 (i.e. 3 individuals x 4 copies = 12 “fake individuals” with P (A=0/L=0)), and 12 copies of the 1 individual with A=1 (i.e. 1 individual x 12 copies = 12 “fake individuals” with P(A=1/L=0)). Thus, there would be an equal number of exposed and unexposed individuals at L=0. We can repeat the same process at L=1, so that those with A=1/L=1 would have an IPTW of 4/3 and those with A=0/L=1 would have an IPTW of 4. If we repeated the same process of multiplying by the relative ratio of 3, we would again end up with 4 copies of each of the three individuals with A=1 (i.e. 3 individuals x 4 copies =12 with P(A=1/L=1)), and 12 copies of the individual with A=0 (i.e. 1 individual x 12 copies = 12 with P(A=1/L=0). In the end, we would thus have, at each level of L, 12 individuals with A=1 and 12 individuals with A=0, and we would have a perfectly balanced distribution of exposure history by covariate history. We then use the weighted “pseudopopulation” to estimate the relationship between exposure A and outcome Y. In this way, inverse probability of weighting addresses a potential imbalance in confounders by exposure history, and thus addresses confounding without introducing the simultaneous confounder and mediator L into the equation estimating the relationship between A and Y.

eAppendix 2: Directed acyclic graph of the hypothesized relationship between neighborhood poverty and alcohol use

Figure A1 presents a directed acyclic graph that specifies the temporal ordering of the associations in the underlying causal model, assuming no unmeasured confounders.

Figure A1. Directed acyclic graph (causal diagram) for the CARDIA study, assuming no unmeasured confounders. $A_0 - A_{15}$ represent the proportion of neighborhood residents living in poverty at each time point, $L_0 - L_{15}$ represent vectors of measured confounders that may be associated with $A_0 - A_{15}$ in time periods 0 to 15.



the notation poposed by Robins et al. [1] and expanded upon by Following Bodnar et al. [2], let A_0 indicate the proportion of neighborhood residents in poverty for a specific respondent's neighborhood at baseline. Let A_2 equal exposure to a proportion of residents in poverty at year 2, and let L_0 represent a vector of measured confounders that may be associated with A_2 (ethnicity/race, age, income up to year 2, education up to year 2, etc.). A is a vector of responses from baseline to year 15 ($\bar{A}$ =(A_0, A_2, A_5, A_7, A_{10}, A_{15}) and L is a vector of responses up to year 10 ($\overline{L}$ =(L_0, L_2, L_5, L_7, L_{10}). Let C_2 denote loss to follow-up (censoring) before year 2 (C_0 = 1 if censored and 0 otherwise). Similarly, let C_5 denote loss to follow-up between year 2 and year 5, C_7 denote loss to follow-up between year 5 and year 7, C_{10} denote loss to follow-up between year 7 and year 10, C_{15} denote loss to follow-up between year 10 and year 15, and C_{20} denote loss to follow-up between year 15 and year 20.

eAppendix 3: Methods used to calculate IPTC weights and sample SAS program to estimate IPTC weights and an MSM model

The treatment weights were estimated as the ratio of conditional probability densities of receiving the exposure history the respondent did indeed receive. To estimate the treatment weights, we used the log of neighborhood poverty as the exposure, as it adequately accounted for the highly skewed nature of the data. Mathematically, the exposure history weight up to time *t*, was defined as:

$$sw_i(t) = \prod_{t=0}^{t} \frac{f(pov_{neigh,i}(t) | \overline{C}_i(t-1)=0, \overline{pov}_{neigh,i}(t-1), V_i)}{f(pov_{neigh,i}(t) | \overline{C}_i(t-1)=0, \overline{pov}_{neigh,i}(t-1), \overline{L}_i(t-1), V_i)}$$

Here $C_i(t-1)$ equals 0 if a respondent participated in the interview and answered the questions about the outcome of interest by time *t-1*, and it equals 1 if a respondent missed an interview or failed to respond to the questions about the outcome of interest by time *t* - 1. $f(pov_{neigh,i}(t) | \overline{C}_i(t-1)=0, \overline{pov}_{neigh,i}(t-1), \overline{L}_i(t-1), V_i)$ stands for the conditional probability density of poverty exposure experienced at time *t* by person *i* with baseline covariate measures V_i, given that he or she has remained uncensored and experienced exposure history $\overline{pov}_{neigh,i}(t-1)$ and time-varying covariates $\overline{L}_i(t-1)$ respectively until *t*-

1. The stabilizing numerator density of exposure,

$f\,(pov_{neigh,i}(t)\,|\,\overline{C_i}(t-1)=0,\overline{pov}_{neigh,i}(t-1),V_i)$ was similarly defined except that only the baseline measures of all covariates (V_i) were used.

To estimate the numerator of the treatment weights, we used a log-normal density with the mean of $\ln(pov_{neigh,t-1,i})$ modeled as the linear regression $\alpha_{0ti}+\alpha_1\ln pov_{neigh,t-1,i}+\alpha_2 t_{ti}+\mathrm{X'V}_i$ and its variance modeled as $\exp(\gamma_{0ti}+\gamma_1\ln pov_{neigh,t-1,i}+\gamma_2 t_{ti}+\mathrm{N'V}_i)$. To estimate the denominator, we used a log-normal model with mean of $\ln(pov_{neigh,t-1,i})$, given by $\eta_{0ti}+\eta_1\ln pov_{neigh,t-1,i}+\eta_2 t_{ti}+\Lambda'\mathrm{L}_{t-1i}+\mathrm{O'V}_i+e_i$ and variance $\exp(\chi_{0ti}+\chi_1\ln pov_{neigh,t-1,i}+\chi_2 t_{ti}+\mathrm{T'L}_{t-1i}+\Psi'\mathrm{V}_i)$. Point estimates of the unknown parameters ($\alpha_{0ti},\alpha_1,\alpha_2,\mathrm{X'},\gamma_{0ti},\gamma_1,\gamma_2,\mathrm{N'},\eta_{0ti},\eta_1,\eta_2,\Lambda',\mathrm{O'},\chi_{0ti},\chi_1,\chi_2,\mathrm{T'},\Psi'$) were thus obtained by pooled linear regression for both mean models, and by pooled log-linear regression of estimated squared-residuals for both variance models. These regression estimates and their predicted values were obtained using SAS PROC GENMOD[3]. A sample program is included below (SAS program). These predicted values were in turn used to construct the treatment weights based on the log-normal density assumption.

To construct the corresponding censoring weights, we defined the censoring indicator *C(t)* to be 1 if a subject missed an interview or failed to respond to the questions about the outcome of interest by time *t* and *C(k) = 0* otherwise. Censoring weights were defined and estimated as in Hernan et al[4]. The sample program below provides code to construct censoring weights.

Weights were not trimmed. A sensitivity analysis with trimmed weights yielded comparable results to those with untrimmed weights.

```
/*TREATMENT WEIGHTS*/
/********Program to calculate the numerator of the IPTW weights
***********************/
proc reg data=cardia_long;
	where cens=0; /*this means for those respondents who have not been censored*/
	model log_povt=log_prepovt a01age1 age25m time a01sex black married widow
divsep eventsscore instrsupprt emotspprt bingea07 a0hs cum_anonprof a03depscore
a03child home1 child1x25 home1x25 prof1x25 a03incom inc1x25; /*baseline
covariates*/
	output out=model1a p=pa0_num r=res0_num; /*here res0_num is the residual*/
run;
/*need to check whether the residual is normally distributed*/
proc univariate data=model1a plot normal;
var res0_num;
run;
/*creating the squared residual from the model output*/
data num_residual;
set model1a;
ressq=res0_num*res0_num;
run;
/*estimating the variance of the residual squared*/
```

```
proc genmod data=num_residual;
	where cens=0;
	model ressq= log_prepovt a01age1 age25m time a01sex black married widow divsep eventsscore instrsupprt emotspprt bingea07 a0hs cum_anonprof a03depscore a03child home1 child1x25 home1x25 prof1x25 a03incom inc1x25/dist=normal link=log;
	output out=model1b p=pa0_num2 ;/*pa0_num2 is the variance, or σ^2 */
run;
proc sort data=model1a;
by id time;
proc sort data=model1b;
by id time;
data num_residual2;
merge model1b model1a;
by id time;
/*here we are using the residual and variance to step-by-step estimate the log normal density function*/
residual=ressq/(2*pa0_num2);                    /*
```

$$\frac{(\ln pov_{neigh,ti} - \ln \hat{pov}_{neigh,ti})^2}{2\sigma^2} \quad */$$

```
exponent=exp(-(residual));                    /*
```

$$\exp-\left(\frac{(\ln pov_{neigh,ti} - \ln \hat{pov}_{neigh,ti})^2}{2\sigma^2}\right)*/$$

den=1/((sqrt(2*3.14159))*(sqrt(pa0_num2))*povt); /* $\frac{1}{\sqrt{2\pi\sigma^2}\, pov_{neigh,ti}}$ */

num_prob=den*exponent;

$$/* \frac{1}{\sqrt{2\pi\sigma^2}\, pov_{neigh,ti}} * \exp - \left(\frac{(\ln pov_{neigh,ti} - \ln \hat{pov}_{neigh,ti})^2}{2\sigma^2} \right) */$$

run;

/*Program to calculate the denominator of the IPTW weights

****************************/

```
proc reg data=cardia_long;
	where cens=0;
	model log_povt=log_prepovt a01age1 age25m time a01sex black married widow
divsep eventsscore instrsupprt emotspprt prebinge pst_inc incx25 pst_hs cum_nonprof
dep child homeown profx25 childx25 homex25 ;
	output out=model1b p=pa0_den r=res0_den;
run;
proc univariate data=model1b plot normal;
var res0_den;
run;
/*creating the squared residual from the model output*/
data den_residual;
set model1b;
ressq=res0_den*res0_den;
run;
```

/*estimating the variance of the residual squared*/

proc genmod data=den_residual;

where cens=0;

model ressq= log_prepovt a01age1 age25m time a01sex black married widow divsep eventsscore instrsupprt emotspprt prebinge pst_inc incx25 pst_hs cum_nonprof dep child homeown profx25 childx25 homex25 /dist=normal link=log;

output out=model1b2 p=pa0_den2 ;/*pa0_num2 is the variance, or σ^2 */

run;

proc sort data=model1b;

by id time;

proc sort data=model1b2;

by id time;

data den_residual2;

merge model1b model1b2;

by id time;

residual=ressq/(2*pa0_den2); /* $\frac{\sqrt{\ln pov_{neigh,ti} - \ln \hat{pov}_{neigh,ti}}}{2\sigma^2}$ */

exponent=exp(-(residual)); /* $\exp-\left(\frac{(\ln pov_{neigh,ti} - \ln \hat{pov}_{neigh,ti})^2}{2\sigma^2}\right)$ */

den=1/((sqrt(2*3.14159))*(sqrt(pa0_den2))*povt); /* $\frac{1}{\sqrt{2\pi\sigma^2}\, pov_{neigh,ti}}$ */

den_prob=den*exponent;

$$/* \frac{1}{\sqrt{2\pi\sigma^2}\, pov_{neigh,ti}} * \exp - \left(\frac{(\ln pov_{neigh,ti} - \ln \hat{pov}_{neigh,ti})^2}{2\sigma^2} \right) */$$

run;

proc univariate data=den_residual2 plot normal;

var stdres den_prob;

run;

/*CENSORING WEIGHTS (IPCW)*/

/*First is the program to estimate the numerator*/

```
proc logistic descending data=cardia_long;
	class a01sex ;
	model cens(event='0')= = log_prepovt a01age1 age25m time a01sex black
married widow divsep eventsscore instrsupprt emotspprt bingea07 a0hs cum_anonprof
a03depscore a03child home1 child1x25 home1x25 prof1x25 a03incom inc1x25;
	output out=model1ca p=pc0_num;
run;
```

/*Second is the program to estimate the denominator*/

```
proc logistic descending data=cardia_long ;
	class a01sex prebinge home;
	model cens(event='0')= log_prepovt a01age1 age25m time a01sex black married
widow divsep eventsscore instrsupprt emotspprt prebinge pst_inc incx25 pst_hs
cum_nonprof dep child homeown profx25 childx25 homex25;
```

```
output out=model1cb p=pc0_den;
run;
proc sort data=num_residual2;
by id time;
run;
proc sort data=den_residual2;
by id time;
run;
proc sort data=model1ca;
by id time;
run;
proc sort data=model1cb;
by id time;
run;
```

/*Here we actually create the final stabilized weights, which are a product of the IPTW and IPCW weights*/

```
data weights;
merge den_residual2 num_residual2 model1b model1ca model1cb;
by id time;
if first.id then do;
k1_0=1;kc1_0=1;
k1_w=1;kc1_w=1;
end;
```

```
	retain k1_0 kc1_0 k1_w kc1_w;
	/*inverse probability of censoring weights*/
	kc1_0=kc1_0*pc0_num;
	kc1_w=kc1_0*pc0_den;
	/*inverse probability of treatment weights*/
				k1_0=k1_0*num_prob;
				k1_w=k1_w*den_prob;


	/*stabilized weights*/
	stabwt=(k1_0*kc1_0)/(k1_w*kc1_w);
	run;
```

/*WEIGHTED MSM MODEL*?

```
 /*Here we use the stabilized weights to estimate an MSM model—that is a weighted
model estimating the marginal relationship between cumulative poverty at t-1 and the
odds of binging in the past month */
proc genmod descending data=weights ;
	class id;
	model binge= a01age1 time cum_povt age25m age a01sex black married widow
divsep
eventsscore instrsupprt emotspprt bingea07 a0hs cum_anonprof a03depscore a03child
home1 child1x25
home1x25 a03incom inc1x25 prof1x25 /link=logit dist=bin type3;
	weight stabwt;
```

```
    repeated subject=id/type=ind;
run;
```

eAppendix 4*:* Testing whether time-varying covariates acted as confounders and mediators

We examined whether the time-varying covariates of interest in our data could be both confounders and mediators in the causal pathway between neighborhood poverty and alcohol use/misuse. This was a necessary precondition for MSMs to be useful. We tested whether: a) key time-varying covariates of interest, low income, non-professional/managerial occupations and low education were longitudinally associated with later neighborhood poverty (i.e. covariates could act as a selector into neighborhood poverty); b) neighborhood poverty predicted key time-varying covariates of interest (i.e. whether time-varying covariates fulfilled the first requirement to be mediators of the neighborhood poverty - binging and neighborhood poverty - frequency of alcohol use relationships); and c) the time-varying covariates were associated with alcohol frequency of use/binging, independently of neighborhood poverty (i.e. the second condition necessary for the covariates to be confounders or mediators of the neighborhood poverty -alcohol use relationship).

In order to test these conditions, we estimated a series of models including: a) repeated measures linear regression models separately estimating the association between lagged income, education and non-professional/managerial status and neighborhood poverty; b) repeated measures marginal logistic regression models separately estimating the association between lagged neighborhood poverty and low income, low education and non-professional/managerial occupational status; c) repeated measures marginal logistic/ negative binomial models estimating the association between lagged neighborhood poverty (estimated separately as cumulative up to *t*-1 and as just poverty at *t*-1), low income, low education, non-professional/managerial occupational status, and alcohol use (frequency of use and binging). Since neighborhood poverty is skewed, ln(neighborhood poverty) was used as the outcome measure in the models that estimated the association between it and lagged income, education and non-professional/managerial status.

An additional precondition for marginal structural models to be useful is that time-varying covariates and the main exposure of interest, in this case neighborhood poverty, actually change over time. We tested this precondition by examining the tracking correlation between reports for the same measure over time; we wanted to see if the magnitude of the correlation between measures of neighborhood poverty at different examination points decreased markedly over time.

Table A1 shows the relationship between income, occupation and education at *t*-1 and ln(neighborhood poverty) at *t*. Model 4 incorporates all three predictors of ln(neighborhood poverty): having a low income, less than high school education and a longer exposure to non-professional/managerial positions were all positively associated with ln(neighborhood poverty).

Table A2 provides estimates of associations in the other direction: the association between lagged neighborhood poverty and the odds of having low income, less than high school education, and a non-professional/managerial occupation. A 20% increase in the proportion of residents in neighborhood poverty was associated with higher odds of having a low income (OR: 1.32; 95% confidence interval (CI): 1.21,1.45), having less than high school education (OR: 1.32; 95% CI: 1.11,1.57) and with having a non-professional or managerial degree (OR: 1.25; 95% CI: 1.14, 1.37).

Table A3 shows the mean, standard deviation and correlations between neighborhood poverty measures across the six examinations of measurement. The mean proportion of poverty in neighborhoods decreased over the study duration, from 24% to 12% of the population. Moreover, the magnitude of the correlation between neighborhood poverty decreased from 0.98 to 0.22 over time, indicating that the concentration of poverty in the neighborhood did change over time.

eTable 1. Parameter estimates and standard errors from mixed linear regression models exploring the relationship between lagged income, occupation and education and dependent variable ln(neighborhood poverty) over time: the CARDIA study, 1985-2001

	M1[1]		M2[2]		M3[3]		M4[4]	
Variable	Parameter	SE	Parameter	SE	Parameter	SE	Parameter	SE
Intercept	-0.82***	0.03	-0.75***	0.04	-0.82***	0.04	-0.88***	0.04
Baseline covariates								
Age (years)	0.002	0.001	0.001	0.001	0.002*	0.001	0.003*	0.001
Time	-0.07***	0.004	-0.07***	0.004	-0.07***	0.004	-0.07***	0.004
Female	-0.01	0.01	-0.005	0.01	-0.01	0.01	-0.01	0.01
Race/ethnicity								
Black	0.20***	0.01	0.21***	0.01	0.21***	0.01	0.19***	0.01
Marital status (reference: never married)								
Married	-0.02~	0.01	-0.02*	0.01	-0.02~	0.01	-0.02~	0.01
Widowed	0.004	0.02	0.01	0.02	0.01	0.02	0.01	0.02
Divorced/ separated	0.01	0.02	0.02	0.02	0.01	0.02	0.01	0.02
Time-varying covariates								
Prior ln(neighborhood poverty)	0.66***	0.01	0.67***	0.01	0.66***	0.01	0.66***	0.01
Prior low income (≤$24,999)[5]	0.07***	0.01					0.06***	0.01
Prior less than HS education			0.07***	0.02			0.05*	0.02
Prior cumulative occupation[6]					0.07***	0.01	0.04*	0.01
Depression (CESD ≥16)[5]	0.04***	0.01	0.04***	0.01	0.04***	0.01	0.04*	0.01
Number of children in the	0.01	0.01	-0.004	0.01	-0.01	0.01	0.005	0.01

household								
Home ownership (home owned is reference category)[5]	-0.01	0.01	-0.05***	0.01	-0.05***	0.01	0.01	0.01

p-values: ~<.10; *<0.05; ***<0.0001

[1] Mixed linear regression model estimating association between low income and ln(neighborhood poverty)

[2] Mixed linear regression model estimating association between low education and ln(neighborhood poverty)

[3] Mixed linear regression model estimating association between cumulative non-professional/managerial occupation and ln(neighborhood poverty)

[4] Mixed linear regression model estimating association between low income, low education, cumulative non-professional/managerial occupation and ln(neighborhood poverty)

[5] Values are interpolated for those examination times when the covariate was not measured

[6] Cumulative occupation is defined as cumulative exposure to non-professional or managerial up to examination *t*-1

eTable 2. Odds ratios and 95% confidence intervals estimating the association between lagged ln(neighborhood poverty) and three time-varying dependent variables: low income (<$25,000), less than high school education and non-professional/managerial occupations: the CARDIA study, 1985-2001

	Low income			Less than high school			Non-professional or managerial occupations		
	M1			M2			M3		
	OR	95% CI		OR	95% CI		OR	95% CI	
Intercept	2.29	1.40	3.73	0.34	0.10	1.24	7.94	5.04	12.50
Baseline covariates									
Age (years)	0.96	0.94	0.97	0.94	0.90	0.98	0.94	0.92	0.95
Time	0.80	0.77	0.82	0.85	0.80	0.91	0.91	0.88	0.93
Female	1.02	0.90	1.14	0.55	0.41	0.72	0.93	0.83	1.04
Race/ethnicity									
Black	1.59	1.39	1.80	1.23	0.88	1.72	2.05	1.82	2.32
Marital status (reference: never married)									
Married	0.44	0.37	0.53	0.96	0.62	1.49	0.89	0.76	1.03
Widowed	0.67	0.53	0.84	0.99	0.55	1.78	0.93	0.76	1.15
Divorced/separated	1.13	0.88	1.43	1.54	0.89	2.65	1.22	0.94	1.58
Time-varying covariates									
Poverty at t-1	4.07	2.60	6.36	3.95	1.66	9.45	3.05	1.94	4.82
Depression (CESD ≥16)	1.64	1.44	1.86	1.94	1.49	2.54	1.42	1.25	1.61
Number of children in the household	1.55	1.36	1.76	2.90	2.13	3.96	1.90	1.70	2.13

Home ownership (reference: owned)	0.37	0.33	0.41	0.37	0.27	0.51	0.77	0.69	0.85

eTable 3. Correlations between neighborhood poverty at different years of measurement: the CARDIA study, 1985-2001

Year of measurement	Neighborhood poverty by year of measurement					
	1985-86	1987-88	1990-91	1992-93	1995-96	2000-01
1985-86	1	0.98	0.67	0.41	0.26	0.22
1987-88	0.98	1	0.75	0.48	0.31	0.25
1990-91	0.67	0.75	1	0.43	0.28	0.19
1992-93	0.41	0.48	0.43	1	0.49	0.3
1995-96	0.26	0.31	0.28	0.49	1	0.27
2000-01	0.22	0.25	0.19	0.3	0.27	1

eAppendix 5: Estimating the association between time-varying covariates and the exposure of interest, neighborhood poverty, after weighting the data by the IPTC

As shown in Table A4 below, we found that the IPTCW-weighted association between each time-varying covariate and the subsequent exposure to neighborhood poverty was null, so that the weights did account for the potential association between time-varying covariates and the exposure of interest.

Table A4. Hierarchical linear mixed model estimating the association between baseline and time-varying characteristics at *t*-1 and poverty at *t*, after weighting by the IPTCW

	Beta estimate	Standard error	P-value
Aged 25 or older at baseline	**0.42**	**0.05**	**<.0001**
Age	**-0.08**	**0.00**	**<.0001**
Sex	-0.01	0.02	0.61
Black race (reference group is white)	**0.50**	**0.02**	**<.0001**
Marital status			
Married	0.05	0.02	0.05
Widowed	0.06	0.03	0.11
Divorced/separated	**0.12**	**0.04**	**0.00**
Life events score	0.00	0.00	0.69
Social support			
Instrumental support	-0.07	0.06	0.27
Emotional support	0.08	0.05	0.15
Binging at *t*-1	0.02	0.02	0.20
Income ($50,000+ is the reference group)			
Category 1 ($0-4,999)	0.07	0.05	0.17
Category 2 ($5-11,999)	0.05	0.04	0.27
Category 3 ($12-15,999)	0.07	0.04	0.13
Category 4 ($16-24,999)	0.03	0.04	0.40
Category 5 ($25-34,999)	0.01	0.04	0.75
Category 6 ($35-49,999)	0.00	0.05	1.00

Category 1 x cohort	0.07	0.07	0.32
Category 2 x cohort	0.12	0.06	0.06
Category 3 x cohort	0.08	0.06	0.18
Category 4 x cohort	0.09	0.05	0.07
Category 5 x cohort	0.08	0.04	0.11
Category 6 x cohort	0.05	0.05	0.32
Less than HS education	0.03	0.03	0.31
Mean number of years in non-professional/managerial occupations	0.06	0.04	0.14
Employment x cohort	**0.04**	**0.02**	**0.02**
Depression (16 or above in CES-D)	-0.02	0.03	0.42
Any children at home	-0.02	0.03	0.55
Children x cohort	0.03	0.05	0.56
Home ownership (reference: owned)	-0.02	0.03	0.60
Home ownership x cohort	-0.07	0.04	0.06

eAppendix 6. Checking whether there is variation in levels of poverty within weight strata, and whether subjects are comparable on covariates in each weight stratum, across different levels of poverty

As shown in Table A5, we checked for the assumption of “common support”, and found that at each weight stratum, there was variation in observed neighborhood poverty, and those subjects exposed to high vs. low levels of neighborhood poverty were comparable on the covariates of interest.

Table A5. Mean subject characteristics by study interview, weight rank, and level of poverty

Time	Weight rank	% of residents in poverty	N	Mean weight	Mean predicted % poverty	Mean age	Mean % female	Mean % black	Mean % in non-prof/ manager jobs	Mean % less than HS education	Mean depressed	Mean % with children	Mean % own home	Mean % with income $0-4,999	Mean % with income $5-11,999	Mean % with income $12-15,999	Mean % with income $16-24,999
1	0	≤10%	106	1.03	0.08	26.70	0.64	0.03	0.56	0.06	0.34	0.32	0.72	0.00	0.04	0.05	0.10
1	0	10-20%	198	1.04	0.13	26.56	0.61	0.10	0.63	0.01	0.54	0.22	0.74	0.01	0.02	0.03	0.13
1	0	20-30%	126	1.03	0.17	26.59	0.67	0.17	0.68	0.03	0.52	0.05	0.67	0.01	0.00	0.00	0.13
1	0	>30%	83	1.01	0.27	25.70	0.63	0.29	0.75	0.03	0.46	0.06	0.52	0.02	0.01	0.03	0.20
1	1	≤10%	232	1.13	0.09	25.62	0.6	0.05	0.56	0.03	0.16	0.10	0.61	0.01	0.02	0.04	0.12
1	1	10-20%	600	1.13	0.13	25.81	0.58	0.24	0.71	0.04	0.26	0.33	0.58	0.03	0.07	0.04	0.20
1	1	20-30%	303	1.13	0.19	25.55	0.53	0.35	0.75	0.03	0.37	0.26	0.44	0.04	0.05	0.06	0.34
1	1	>30%	172	1.13	0.27	24.75	0.65	0.58	0.79	0.02	0.48	0.22	0.43	0.04	0.06	0.06	0.31
1	2	≤10%	131	1.23	0.09	24.73	0.52	0.11	0.74	0.08	0.03	0.09	0.43	0.02	0.01	0.05	0.15
1	2	10-20%	517	1.23	0.14	24.72	0.52	0.47	0.80	0.07	0.18	0.32	0.39	0.06	0.10	0.09	0.16
1	2	20-30%	372	1.24	0.2	24.95	0.57	0.58	0.84	0.10	0.25	0.37	0.39	0.05	0.13	0.12	0.22
1	2	>30%	248	1.24	0.29	24.34	0.62	0.72	0.88	0.11	0.38	0.39	0.36	0.08	0.13	0.11	0.23
1	3	≤10%	39	1.35	0.11	24.24	0.4	0.53	0.75	0.07	0.09	0.10	0.46	0.02	0.06	0.04	0.18
1	3	10-20%	328	1.35	0.16	23.48	0.49	0.79	0.85	0.15	0.07	0.31	0.31	0.07	0.11	0.12	0.16
1	3	20-30%	313	1.36	0.21	24.20	0.53	0.77	0.85	0.13	0.12	0.45	0.33	0.08	0.13	0.14	0.15
1	3	>30%	285	1.37	0.3	24.22	0.55	0.84	0.87	0.15	0.15	0.51	0.35	0.10	0.18	0.10	0.15
1	4	≤10%	8	1.57	0.1	24.90	0.2	0.98	0.75	0.13	0.00	0.23	0.28	0.00	0.00	0.00	0.13
1	4	10-20%	78	1.54	0.17	22.63	0.34	0.97	0.80	0.39	0.04	0.26	0.26	0.16	0.06	0.11	0.15
1	4	20-30%	87	1.57	0.23	22.55	0.35	0.94	0.78	0.32	0.01	0.46	0.33	0.15	0.09	0.18	0.04
1	4	>30%	102	1.56	0.3	23.07	0.38	0.91	0.82	0.28	0.04	0.55	0.28	0.19	0.12	0.20	0.04
2	0	≤10%	213	1.00	0.08	28.52	0.6	0.06	0.51	0.02	0.25	0.31	0.74	0.00	0.04	0.03	0.09
2	0	10-20%	289	1.00	0.12	28.27	0.58	0.20	0.62	0.01	0.40	0.30	0.65	0.02	0.04	0.03	0.13
2	0	20-30%	176	0.99	0.17	27.79	0.67	0.36	0.72	0.04	0.50	0.22	0.57	0.01	0.01	0.02	0.22
2	0	>30%	98	0.96	0.23	26.95	0.53	0.51	0.73	0.03	0.39	0.14	0.47	0.04	0.03	0.04	0.22
2	1	≤10%	213	1.13	0.08	27.51	0.57	0.09	0.61	0.02	0.17	0.21	0.62	0.01	0.03	0.06	0.15
2	1	10-20%	324	1.13	0.13	27.35	0.54	0.29	0.68	0.03	0.26	0.36	0.51	0.04	0.09	0.05	0.23
2	1	20-30%	186	1.13	0.18	27.21	0.63	0.50	0.77	0.04	0.42	0.33	0.44	0.07	0.06	0.07	0.31

2	1	>30%	82	1.13	0.24	26.73	0.59	0.69	0.80	0.05	0.37	0.38	0.34	0.04	0.09	0.08	0.30
2	2	≤10%	169	1.24	0.09	27.16	0.53	0.17	0.69	0.03	0.11	0.19	0.48	0.02	0.04	0.06	0.16
2	2	10-20%	309	1.24	0.13	26.78	0.52	0.41	0.75	0.04	0.19	0.32	0.40	0.04	0.12	0.07	0.21
2	2	20-30%	217	1.24	0.18	26.94	0.62	0.63	0.79	0.06	0.28	0.45	0.42	0.08	0.10	0.09	0.29
2	2	>30%	97	1.24	0.25	26.08	0.56	0.82	0.88	0.06	0.42	0.45	0.36	0.09	0.12	0.10	0.30
2	3	≤10%	121	1.36	0.09	26.56	0.54	0.30	0.75	0.06	0.11	0.25	0.45	0.02	0.05	0.09	0.16
2	3	10-20%	333	1.37	0.14	26.42	0.53	0.55	0.78	0.04	0.14	0.32	0.32	0.05	0.08	0.10	0.17
2	3	20-30%	260	1.37	0.19	26.52	0.55	0.72	0.79	0.09	0.22	0.44	0.35	0.09	0.15	0.13	0.16
2	3	>30%	143	1.37	0.25	26.26	0.57	0.79	0.85	0.08	0.26	0.53	0.34	0.08	0.24	0.09	0.19
2	4	≤10%	55	1.61	0.09	26.79	0.42	0.59	0.78	0.07	0.08	0.17	0.40	0.04	0.03	0.07	0.19
2	4	10-20%	168	1.57	0.14	25.85	0.49	0.71	0.79	0.11	0.06	0.28	0.38	0.07	0.06	0.12	0.13
2	4	20-30%	241	1.61	0.19	25.96	0.48	0.81	0.80	0.15	0.10	0.54	0.35	0.09	0.13	0.19	0.05
2	4	>30%	177	1.63	0.25	26.27	0.52	0.82	0.83	0.14	0.12	0.59	0.33	0.14	0.13	0.14	0.08
3	0	≤10%	457	0.92	0.08	31.00	0.59	0.15	0.58	0.01	0.25	0.52	0.65	0.01	0.04	0.06	0.09
3	0	10-20%	158	0.96	0.12	30.84	0.64	0.27	0.62	0.00	0.43	0.34	0.58	0.02	0.03	0.01	0.20
3	0	20-30%	97	0.96	0.15	30.52	0.65	0.53	0.75	0.03	0.49	0.50	0.47	0.02	0.03	0.06	0.24
3	0	>30%	79	0.92	0.17	30.05	0.59	0.58	0.81	0.05	0.34	0.37	0.46	0.05	0.05	0.07	0.26
3	1	≤10%	324	1.13	0.09	30.50	0.55	0.15	0.63	0.02	0.24	0.41	0.57	0.03	0.07	0.05	0.16
3	1	10-20%	138	1.13	0.12	30.53	0.62	0.40	0.69	0.01	0.36	0.40	0.43	0.06	0.09	0.04	0.26
3	1	20-30%	83	1.13	0.15	30.12	0.66	0.68	0.80	0.03	0.41	0.55	0.43	0.05	0.04	0.12	0.42
3	1	>30%	51	1.13	0.17	29.75	0.53	0.71	0.84	0.07	0.36	0.50	0.36	0.05	0.15	0.05	0.30
3	2	≤10%	303	1.24	0.09	29.95	0.54	0.22	0.65	0.03	0.16	0.42	0.47	0.04	0.06	0.06	0.18
3	2	10-20%	139	1.23	0.13	29.77	0.55	0.50	0.73	0.03	0.24	0.47	0.35	0.05	0.12	0.06	0.31
3	2	20-30%	100	1.24	0.15	30.17	0.62	0.80	0.85	0.05	0.32	0.63	0.42	0.06	0.15	0.08	0.38
3	2	>30%	59	1.24	0.18	28.90	0.58	0.80	0.90	0.06	0.37	0.55	0.35	0.08	0.14	0.10	0.34
3	3	≤10%	314	1.37	0.1	29.80	0.51	0.37	0.69	0.03	0.14	0.42	0.42	0.05	0.05	0.08	0.14
3	3	10-20%	167	1.37	0.14	29.49	0.56	0.63	0.76	0.05	0.20	0.53	0.32	0.05	0.13	0.13	0.19
3	3	20-30%	115	1.37	0.16	29.76	0.54	0.83	0.84	0.06	0.23	0.65	0.36	0.07	0.17	0.11	0.29
3	3	>30%	74	1.37	0.19	28.64	0.52	0.91	0.88	0.10	0.30	0.62	0.25	0.16	0.24	0.06	0.25
3	4	≤10%	317	1.82	0.11	29.42	0.52	0.52	0.72	0.06	0.14	0.41	0.44	0.05	0.04	0.08	0.18
3	4	10-20%	209	1.64	0.14	29.33	0.57	0.62	0.76	0.05	0.11	0.51	0.39	0.07	0.10	0.20	0.06
3	4	20-30%	163	1.69	0.18	29.33	0.48	0.89	0.83	0.11	0.15	0.71	0.39	0.08	0.15	0.16	0.08
3	4	>30%	129	1.69	0.2	28.84	0.49	0.90	0.86	0.15	0.19	0.74	0.30	0.14	0.17	0.14	0.16
4	0	≤10%	503	0.91	0.06	33.05	0.58	0.19	0.60	0.01	0.22	0.56	0.68	0.01	0.06	0.06	0.08
4	0	10-20%	139	0.96	0.09	32.79	0.58	0.35	0.61	0.01	0.33	0.47	0.66	0.02	0.03	0.02	0.15
4	0	20-30%	55	0.95	0.13	32.74	0.68	0.53	0.67	0.00	0.35	0.49	0.57	0.05	0.00	0.03	0.18
4	0	>30%	46	0.88	0.15	32.54	0.74	0.49	0.78	0.03	0.43	0.48	0.57	0.07	0.09	0.07	0.29
4	1	≤10%	277	1.13	0.06	32.48	0.57	0.21	0.63	0.02	0.22	0.51	0.62	0.02	0.07	0.03	0.13

4	1	10-20%	91	1.13	0.09	32.06	0.56	0.39	0.70	0.02	0.32	0.47	0.52	0.05	0.06	0.04	0.20
4	1	20-30%	42	1.13	0.13	32.40	0.62	0.62	0.78	0.04	0.27	0.54	0.49	0.08	0.09	0.11	0.22
4	1	>30%	23	1.13	0.15	32.23	0.59	0.62	0.85	0.00	0.38	0.54	0.41	0.08	0.18	0.09	0.21
4	2	≤10%	264	1.24	0.07	32.16	0.54	0.31	0.66	0.03	0.19	0.52	0.57	0.02	0.08	0.06	0.16
4	2	10-20%	110	1.24	0.1	31.80	0.61	0.45	0.71	0.01	0.26	0.45	0.44	0.04	0.10	0.07	0.18
4	2	20-30%	52	1.24	0.15	32.06	0.61	0.74	0.80	0.03	0.25	0.57	0.49	0.06	0.11	0.06	0.22
4	2	>30%	24	1.24	0.17	32.37	0.55	0.76	0.85	0.10	0.35	0.63	0.46	0.16	0.15	0.06	0.18
4	3	≤10%	305	1.37	0.07	31.89	0.52	0.40	0.69	0.03	0.18	0.56	0.49	0.03	0.09	0.07	0.16
4	3	10-20%	132	1.37	0.11	31.32	0.58	0.60	0.70	0.02	0.20	0.48	0.42	0.04	0.10	0.09	0.16
4	3	20-30%	84	1.37	0.14	32.03	0.53	0.79	0.81	0.02	0.18	0.66	0.46	0.04	0.09	0.08	0.17
4	3	>30%	33	1.37	0.17	31.55	0.53	0.78	0.83	0.12	0.40	0.63	0.36	0.06	0.21	0.11	0.14
4	4	≤10%	430	1.84	0.07	31.52	0.53	0.50	0.72	0.05	0.14	0.54	0.53	0.04	0.09	0.05	0.14
4	4	10-20%	232	1.73	0.12	31.46	0.56	0.68	0.75	0.04	0.12	0.60	0.47	0.05	0.09	0.13	0.15
4	4	20-30%	138	1.72	0.15	31.64	0.54	0.82	0.80	0.04	0.17	0.64	0.47	0.09	0.12	0.10	0.11
4	4	>30%	84	1.92	0.17	30.69	0.46	0.82	0.83	0.14	0.24	0.67	0.38	0.13	0.09	0.15	0.10
5	0	≤10%	454	0.91	0.05	36.04	0.55	0.16	0.55	0.01	0.18	0.64	0.77	0.02	0.04	0.02	0.06
5	0	10-20%	134	0.91	0.07	35.75	0.69	0.39	0.68	0.01	0.33	0.57	0.61	0.02	0.07	0.06	0.16
5	0	20-30%	50	0.86	0.08	36.23	0.65	0.50	0.73	0.03	0.37	0.63	0.52	0.09	0.06	0.11	0.08
5	0	>30%	29	0.90	0.09	36.08	0.45	0.38	0.65	0.06	0.34	0.50	0.45	0.10	0.12	0.11	0.17
5	1	≤10%	224	1.13	0.05	35.43	0.56	0.18	0.61	0.01	0.18	0.57	0.71	0.01	0.03	0.03	0.09
5	1	10-20%	82	1.13	0.07	35.38	0.63	0.47	0.66	0.01	0.33	0.55	0.61	0.04	0.09	0.04	0.10
5	1	20-30%	30	1.13	0.09	35.90	0.62	0.69	0.76	0.03	0.39	0.79	0.59	0.05	0.15	0.05	0.08
5	1	>30%	14	1.13	0.11	34.37	0.56	0.74	0.86	0.00	0.43	0.66	0.38	0.04	0.05	0.22	0.33
5	2	≤10%	231	1.24	0.06	35.25	0.57	0.30	0.61	0.02	0.19	0.59	0.63	0.02	0.04	0.03	0.12
5	2	10-20%	87	1.24	0.08	34.80	0.59	0.53	0.72	0.03	0.29	0.53	0.52	0.04	0.07	0.06	0.14
5	2	20-30%	40	1.23	0.1	35.74	0.65	0.71	0.81	0.03	0.33	0.76	0.63	0.06	0.12	0.07	0.08
5	2	>30%	22	1.24	0.1	35.39	0.63	0.78	0.78	0.01	0.44	0.61	0.34	0.08	0.22	0.11	0.20
5	3	≤10%	263	1.37	0.06	34.98	0.53	0.33	0.64	0.03	0.16	0.59	0.62	0.03	0.03	0.04	0.11
5	3	10-20%	99	1.37	0.08	34.74	0.56	0.57	0.73	0.04	0.24	0.62	0.50	0.04	0.05	0.06	0.15
5	3	20-30%	55	1.37	0.1	35.20	0.6	0.80	0.80	0.06	0.25	0.69	0.56	0.09	0.09	0.09	0.12
5	3	>30%	27	1.37	0.1	34.79	0.63	0.66	0.85	0.03	0.42	0.67	0.37	0.08	0.13	0.08	0.14
5	4	≤10%	425	1.83	0.06	34.60	0.52	0.44	0.67	0.04	0.12	0.65	0.62	0.04	0.04	0.05	0.11
5	4	10-20%	250	1.86	0.09	34.44	0.56	0.69	0.76	0.03	0.13	0.62	0.47	0.04	0.06	0.11	0.12
5	4	20-30%	117	1.79	0.1	34.40	0.51	0.82	0.82	0.06	0.19	0.64	0.42	0.06	0.11	0.12	0.18
5	4	>30%	81	1.91	0.11	34.31	0.6	0.81	0.83	0.07	0.17	0.80	0.39	0.06	0.15	0.11	0.13

REFERENCES

1. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11(5):550-560.

2. Bodnar LM, Davidian M, Siega-Riz AM, Tsiatis AA. Marginal Structural Models for Analyzing Causal Effects of Time-dependent Treatments: An Application in Perinatal Epidemiology. American Journal of Epidemiology 2004;159(10):926-934.
3. SAS. SAS/STAT User's Guide. Cary, NC, 1999.
4. Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology 2000;11(5):561-570.